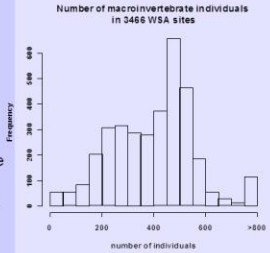
The Influence of Reducing Full Macroinvertebrate Sample Data to a Common Fixed 300 Individual count on Assessments of Stream Quality

Jeannie Sifneos and Alan Herlihy, Departments of Statistics and Fisheries and Wildlife, Oregon State University

The Problem:

Scientists need a fixed count to normalize count effort when calculating stream assessment metrics. The macroinvertebrate data rarely return from the lab with the exact fixed count. To alleviate this problem, researchers take a random sample of individuals so that all sites have the same number of bugs used to calculate metrics. The random sample can introduce variability into the final metric calculation. Our objective was to quantify the variability due to the random sampling on 4 bioassessment metrics of stream quality: a multimetric index (MMI) with range 0.35-96, total taxa RICHNESS with range 9-82, Ephemeroptera, Plecoptera and Trichoptera (EPT) RICHNESS with range 0-37 and predictive model observed/expected taxa ratio (O/E) with range 0.11-1.63.



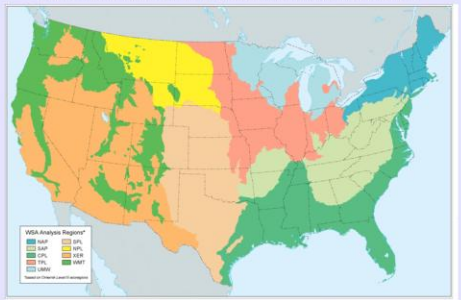
Data set: We used sites with 300 or more individuals from the 3466 sites in the EPA's National Wadeable Stream Assessment.

Methods:

- 1000 random samples of 300 individuals each were drawn from each site's complete macroinvertebrate abundance data list.
- For each sample, the 6 metrics (e.g., % EPT taxa) used to calculate the MMI score and TOTAL RICHNESS, EPT RICHNESS and O/E were calculated.
- The coefficient of variation and standard deviation of the 1000 resamples were calculated for each metric.
- The MMI is based on scoring developed for each of the 9 aggregated WSA ecoregions (Stoddard et al., 2008). The O/E models were developed for 3 broader ecoregions (Appalachian Mountains (SAP, NAP), the West (XER, WMT) and the Plains (CPL, NPL, SPL, TPL, UMW) as in Yuan et al. (2008)).

Aggregated Ecoregion	Number of sites	Number of reference sites
Southern Plains (SPL)	97	41
Western Mountains (WMT)	907	464
Xeric (XER)	296	168
Coastal Plain (CPL)	128	55
Northern Plain (NPL)	118	25
Temperate Plains (TPL)	220	74
Northern Appalachians (NAP)	166	125
Southern Appalachians (SAP)	427	291
Upper Midwest (UMW)	110	25
Total	2469	1268

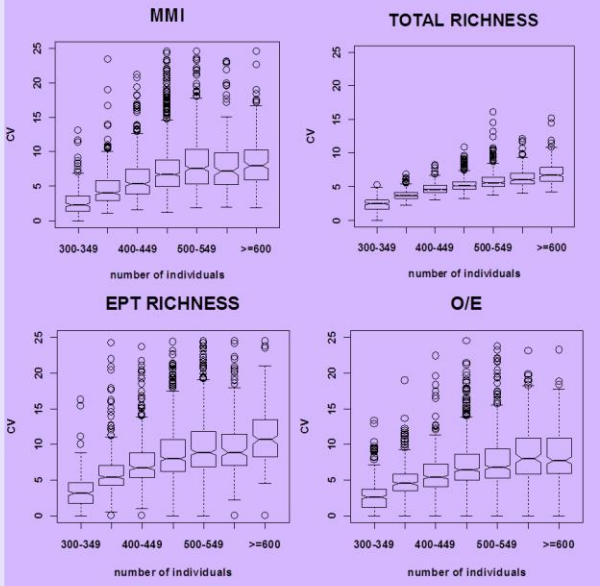
EPA's WADEABLE STREAM ASSESSMENT ECOREGIONS



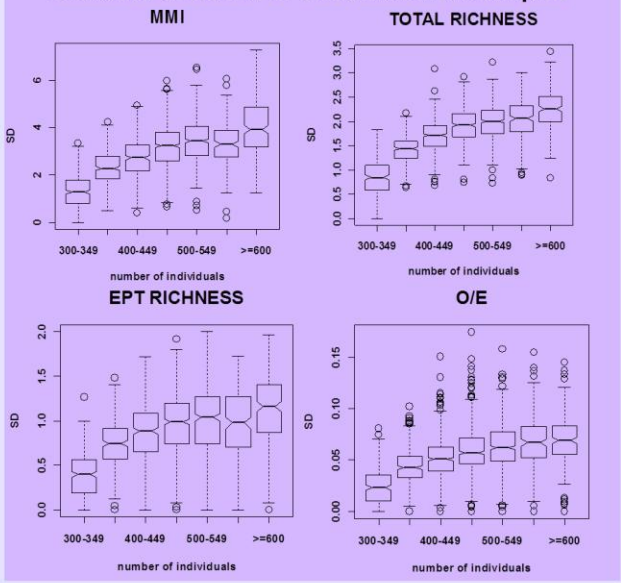
References: Yuan, L., Hawkins, C., and J. Van Sickle. 2008. Effects of regionalization decisions on an O/E index for the US national assessment. J.N. Am. Benthol. Soc. 27:892-905.

Stoddard, J., Herlihy, A., Peck, D., Hughes, R., and T. Whittier. 2008. A process for creating multimetric indices for large-scale aquatic surveys. J.N. Am. Benthol. Soc. 27:878-891.

Coefficient of variation of 1000 random resamples



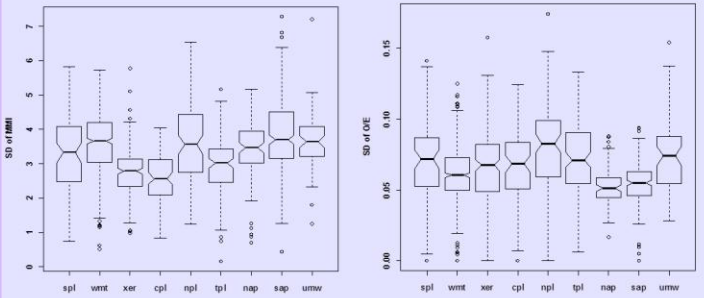
Standard deviation of 1000 random resamples



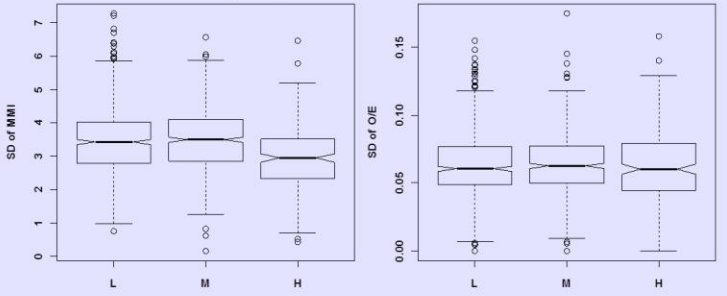
Note: Outliers out of range of y axis not shown. Whiskers go to observations that are ≤ 1.5 * IQR of data.

MEDIAN (and IQR) OF CV AND SD OF 1000 RESAMPLES FOR SITES WITH 450 OR MORE INDIVIDUALS		
	CV	SD
MMI	7.3 (5.1)	3.4 (1.3)
TOTAL RICHNESS	5.5 (1.5)	2.0 (0.52)
EPT RICHNESS	8.9 (5.6)	1.0 (0.51)
O/E	6.8 (4.2)	.062 (.029)

Resampling SD by Ecoregion for MMI and O/E (sites with ≥ 450 individuals)



Resampling SD by least (L), moderately (M) and highly (H) disturbed sites (with ≥ 450 individuals)



Conclusions:

- Resampling does introduce variability into stream assessment metrics.
- Based on 1000 random resamplings of each WSA benthic sample to a fixed 300 count, the median standard deviation for MMI was 3.4 (out of 0-100) and for O/E was .062 (out of 0-1).
- For most of the metrics, the variability increased as the difference between the total number of individuals and 300 increased but did level off above 500 individuals.
- When calculating the MMI, richness metrics introduce more variability than metrics based on percent individuals.
- The resampling variability did not have a strong relationship with ecoregion and was similar among least, moderately and highly disturbed sites.
- We fit one O/E model and then resampled the benthic data 1000 times to look at the variability in "O". It might be interesting to also resample the reference data repeatedly to also assess its effect on the variability in "E" and the final O/E score.